# SEARCH REQUEST FORM

Requestor's Name: S. Unger

Serial Number: -08/848,439

Date: 7/6/98

Phone: 305-2181

Art Unit: 1642

## Search Topic:

Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevent citations, authors, keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevent claim(s).

· SEQ ID #2

SEQ ID #3

Please Compare SEQID #1 to Reverse Transcribed

SEQ ID #2

Please Compare Seq ID# 1 To Reverse Transcribed

SEQ ID #3

Thanks

Susan Yo

1998 JUL -7 AM 6:37

---

## STAFF USE ONLY

Date completed: 7-10-98

Searcher: mazil

Terminal time: 10

Elapsed time: prep 1

CPU time:

Total time: 11

Number of Searches: 1

Number of Databases: 1

**Search Site**
_____ STIC
__/__ CM-1
_____ Pre-S

**Type of Search**
__/__ N.A. Sequence
_____ A.A. Sequence
_____ Structure
_____ Bibliographic

**Vendors**
__/__ IG EG
_____ STN
_____ Dialog
_____ APS
_____ Geninfo
_____ SDC
_____ DARC/Questel
_____ Other

```
> O <
O| |O   IntelliGenetics
> O <

FastDB - Fast Pairwise Comparison of Sequences
Release 5.4

Query sequence being compared:US-08-848-439-1 (1-2027)
Number of sequences searched:    2
Number of scores above cutoff:   2

Results file us-08-848-439-1.res made by spencer on Fri 10 Jul 98 15:04:54-PDT.

Results of the initial comparison of US-08-848-439-1 (1-2027) with:
File : translate.seq
```

```
        100-



         50-



         10-


          5-



          0-

SCORE   0    79   158   238   317   396   475   555   634   713
STDEV            -9   -8   -7   -6   -5   -4   -3   -2   -1    0
                                                    *       *
```

```
                PARAMETERS

Similarity matrix    Unitary
Mismatch penalty     1
Gap penalty          1.00        K-tuple            4
Gap size penalty     0.33        Joining penalty    30
Cutoff score         1           Window size        32
Randomization group  0


                SEARCH STATISTICS

Scores:      Mean      Median    Standard Deviation
             682       653       43.13

Times:       CPU                 Total Elapsed
             00:00:00.00         00:00:00.00

Number of residues:                1710
Number of sequences searched:         2
Number of scores above cutoff:        2
```

The scores below are sorted by initial score.
Significance is calculated based on initial score.

A 100% identical sequence to the query sequence was not found.

The list of best scores is:

| Sequence Name | Description | Init. Score | Length | Opt. Score | Sig. | Frame |
|---|---|---|---|---|---|---|
| 1. US-08-848-439-2 | Sequence 2, Application US | 885 | 713 | 734 | 0.72 | 0 |
| 2. US-08-848-439-3 | Sequence 3, Application US | 825 | 652 | 688 | -0.70 | 0 |

```
1. US-08-848-439-1 (1-2027)
   US-08-848-439-2 Sequence 2, Application US/08848439
```

```
Sequence 2, Application US/08848439
GENERAL INFORMATION:
APPLICANT: LaVALLIE, EDWARD
APPLICANT: RACIE, LISA
TITLE OF INVENTION: HUMAN SDF-5 PROTEIN AND COMPOSITIONS
NUMBER OF SEQUENCES: 3
CORRESPONDENCE ADDRESS:
    ADDRESSEE: GENETICS INSTITUTE, INC.
    STREET: 87 CAMBRIDGEPARK DRIVE
    CITY: CAMBRIDGE
    STATE: MA
    COUNTRY: USA
    ZIP: 02140
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: Patentin Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/848,439
    FILING DATE:
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: LAZAR, STEVEN R.
    REGISTRATION NUMBER: 32,618
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (617) 498-8260
    TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
    LENGTH: 295 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
MOLECULE TYPE: protein

Reverse Translation from the peptide US-08-848-439-2.
Note: the original peptide US-08-848-439-2 contained at least one of
    the residues Arg, Ile, Leu, or Ser.  The nucleic acid sequence
    thus contains ambiguous bases which may translate into amino acids
    other than the original amino acids.
```

| | | | | | |
|---|---|---|---|---|---|
| Initial Score | = | 713 | Optimized Score | = | 734 | Significance | = | 0.72 |
| Residue Identity | = | 58% | Matches | = | 734 | Mismatches | = | 151 |
| Gaps | = | 0 | Conservative Substitutions | = | 0 |

```
       210        220        230        240        250      X 260        270
GCCCTTCCCGGGCTCCGGCTCCCCTCGCCCCTCGGGGGTCGCGCAGGGCGCCCTTGCTCGC
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
                                        ATGYTNCARGGNCCNGGNWNSNY
                                        X                  10                20

       280        290        300        310        320        330        340
TGCTGGCTGCTGCTTCCTCGCCTCGGCCACTCGCTGCCTGGGCTGCGCGGGCTTCGCAGCCCG
||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
TNTNTNYTNTNTTYTNGCNWSNCAYTGYTGYTNGGNWNSNGCNMGNGGNYTNTNTTYTNTTYGGNCARCCNG
    30        40        50        60        70        80        90

       350        360        370        380        390        400        410        420
ACTTCTCCTACAAGCGCAGCAATTGCAAGCCCATCCCCGGCCCAACCTGCCAGCGCTGCCACGGCGGCCATCGGAATACC
|||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
AYTTYWSNTAYAARMGNWSNAAYTGYAARCCNATNCCNGGNCCNAAYTNCARYTNTNGYCAYGGNGGNATNGARTAYC
   100        110        120        130        140        150        160
```

GENERAL INFORMATION:
    APPLICANT: LeVALLIE, EDWARD
    APPLICANT: RACIE, LISA
    TITLE OF INVENTION: HUMAN SDF-5 PROTEIN AND COMPOSITIONS
    NUMBER OF SEQUENCES: 3

2. US-08-848-439-1 (1-2027) : Sequence 3, Application US/08848439
US-08-848-439-3 Sequence 3, Application US/08848439

CORRESPONDENCE ADDRESS:
    ADDRESSEE: GENETICS INSTITUTE, INC.
    STREET: 87 CAMBRIDGEPARK DRIVE
    CITY: CAMBRIDGE
    STATE: MA
    COUNTRY: USA
    ZIP: 02140
COMPUTER READABLE FORM:
    MEDIUM TYPE: Floppy disk
    COMPUTER: IBM PC compatible
    OPERATING SYSTEM: PC-DOS/MS-DOS
    SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
    APPLICATION NUMBER: US/08/848,439
    FILING DATE:
    CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
    NAME: LAZAR, STEVEN R.
    REGISTRATION NUMBER: 32,618
TELECOMMUNICATION INFORMATION:
    TELEPHONE: (617) 498-8260
    TELEFAX: (617) 876-5851
INFORMATION FOR SEQ ID NO: 3:
    SEQUENCE CHARACTERISTICS:
    LENGTH: 275 amino acids
    TYPE: amino acid
    STRANDEDNESS: single
    TOPOLOGY: linear
    MOLECULE TYPE: protein

Reverse Translation from the peptide US-08-848-439-3.
Note: the original peptide US-08-848-439-3 contained at least one of
the residues Arg, Ile, Leu, or Ser. The nucleic acid sequence
thus contains ambiguous bases which may translate into amino acids
other than the original amino acids.

| | | | | |
|---|---|---|---|---|
| Initial Score | = | 652 | Optimized Score | = 688 | Significance | = -0.70 |
| Residue Identity | = | 59% | Matches | = 688 | Mismatches | = 137 |
| Gaps | = | 0 | Conservative Substitutions | = 0 | |

```
      320       330       340       350       360       370       380

700       710       720       730       740       750       760
ACAACGACCTTTGCATCCCCTCGCACCAGCACCACCTTCTGCCAGCCACCGAGCAAGCTTCCAAAGGTAT
||||||||| ||||||||||||| |||  ||| |  |  |||     |  | | | ||||  ||||||||
AYAAYGAYYTNTGYATNCCNYTNGCNWSNWSNGAYCAYTTNYTNCCNACNGARGCNCCNAARGTNT
          390       400       410       420       430       440       450

770       780       790       800       810       820       830       840
GTGAACCCTGCAAAAATAAAAAATGATGATGCAACGACACTAATTGGAAAACGCTTTGTAAAAATGATTTTGCAC
||||| ||| ||  | || | || |||||| | |  | |     ||| |  | |  | |||  | || || |
GYGARGCNTGYAARAAYAARAAYGAYGAYGCNACNAAYTNATNGGNAARAYNYTNGTNAARAAYGAYTTYGCNY
          460       470       480       490       500       510       520

850       860       870       880       890       900       910
TGAAAATAAAAGTGAAGGAGAATAAACCTACACCGAGGATACCAAAAATCATCCTGGAGACCAAGAGGCAAGA
|| | || |  |  |||| | || | |  | || ||| | |  |   |||   |  ||| ||  |  | |
TNAAARATNAARGTNAARGARAAYANCNTAYAYAT NAARAYMGNGAYACNAAARATNATNYTNGARACNAARWSNAARA
          530       540       550       560       570       580       590

920       930       940       950       960       970       980
CCATTTACAACCTGAACGGTGTGTCCGAAAAGGGACCTGAAGAAATCCGGTCGTGTGGCTCAAAGACAGCTTGC
| || |  | | || || | |  |  || |  | |   |    | | | | | |     |  |   |  |  |
CNATNTAYAARYTNAAYGGNGTNWSNGARMGNGCAYYTNAARARWSNGTNYNYNTGGYTNAARGAYWSNYTNC
          600       610       620       630       640       650       660

990       1000       1010       1020       1030       1040       1050
AGTGCACCTGTGAGGAGGATGAAACGACACATCAACGCGCCCTATCCTGGTCATGGGACAGAAAACAGGGTGGGGAGC
|||| | |  | || ||||| || | || | || | | |  |  | |  | | | || |   | |   | |  |
ARTGYACNTGYGARGARRATGAAYGCNCNTAYYTNGTNAATGGGNCARAARCARCARGGNGARY
          680       690       700       710       720       730       740

1060       1070       1080       1090       1100       1110       1120
TGGTGATCACCTCGGTGAAGCGGTGGGCAGAAGGGGCCAGAGAGAGTTCAAGCCGCATCTCCCGCAGCATCCGCA
| ||  |  | | | | || |  | ||| |  ||| | | ||| ||| | | || | | |  | ||| | ||
TNGTNATNACNWSNGTNAARMGNTGGNCARARRGGNCARMGNGARTTYAARMGNMGNATNWSNMGNWSNATNMGNA
          750       760       770       780       790       800       810

1130       1140       1150       1160       1170       1180       1190
AGCTGCAGTGCTAGTCCGGCATCCTGATGGCTCCGACAGGCGCCTCCAGAGCACGGCTG
||| || ||||||| ||
ARYTNCARTGY
          820
          x
```